

Targeting Well-Balanced Solutions in Multi-Objective Bayesian Optimization under a Restricted Budget

D. Gaudrie¹, R. Le Riche², V. Picheny³, B. Eaux¹, V. Herbert¹

¹ Groupe PSA

² CNRS LIMOS, École Nationale Supérieure des Mines de Saint-Étienne

³ Institut National de la Recherche Agronomique, MIAT

Abstract. Multi-objective optimization aims at finding trade-off solutions to conflicting objectives. These constitute the Pareto optimal set. In the context of expensive-to-evaluate functions, it is impossible and often non-informative to look for the entire set. As an end-user would typically prefer solutions with equilibrated trade-offs between the objectives, we define a Pareto front center. We then modify the Bayesian multi-objective optimization algorithm which uses Gaussian Processes to maximize the expected hypervolume improvement, to restrict the search to the Pareto front center. The cumulated effects of the Gaussian Processes and the center targeting strategy lead to a particularly efficient convergence to a critical part of the Pareto set.

Keywords: Gaussian Processes · Parsimonious Optimization · Computer Experiments · Preference-Based Optimization

1 Introduction

Multi-objective optimization aims at minimizing m objectives simultaneously: $\min_{\mathbf{x} \in X \subset \mathbb{R}^d} (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$. As these objectives are generally competing, optimal trade-off solutions $\mathbf{x}_1^*, \dots, \mathbf{x}_q^*$ known as the Pareto optimal set \mathcal{P}_X are sought. These solutions are *non-dominated* (ND): it is not possible to improve one objective without worsening another ($\forall i = 1, \dots, q, \nexists \mathbf{z} \preceq \mathbf{x}_i^*$). The image of \mathcal{P}_X in the objective space is called the Pareto front, $\mathcal{P}_Y = \{f(\mathbf{x}), \mathbf{x} \in \mathcal{P}_X\}$. The Ideal and the Nadir points bound the Pareto front and are defined respectively as $\mathbf{I} = (\min_{\mathbf{y} \in \mathcal{P}_Y} y_1, \dots, \min_{\mathbf{y} \in \mathcal{P}_Y} y_m)$ and $\mathbf{N} = (\max_{\mathbf{y} \in \mathcal{P}_Y} y_1, \dots, \max_{\mathbf{y} \in \mathcal{P}_Y} y_m)$.

Multi-objective optimization algorithms aim at constructing the best approximation to \mathcal{P}_Y , called the *empirical* Pareto front $\widehat{\mathcal{P}}_Y$ which is made of non-dominated observations. At the end of the search, $\widehat{\mathcal{P}}_Y$ is delivered to a Decision Maker (DM) who will choose the solution he/she prefers.

However, when dealing with expensive computer codes, only a few designs \mathbf{x} can be evaluated. In Bayesian optimization, a surrogate is built for each objective upon all past evaluations using Gaussian Process (GP) regression [1].

Information given by these metamodels is used in order to sequentially evaluate new promising inputs with the aim of reaching the Pareto front. As the latter encompasses a large number of solutions when *many* objectives are considered, it may be impossible to compute an accurate approximation within the restricted computational budget. That approximation being used afterwards by a DM, it may anyway be irrelevant to provide the whole Pareto front because it will contain many uninteresting solutions from the DM’s point of view.

Without additional knowledge about the preferences of the DM, we argue that “well-balanced” solutions, in a sense that will be defined hereafter as the central part of the Pareto front, are the most interesting ones. Therefore, we show how to estimate the center of the Pareto front and how classical infill criteria used in Bayesian optimization can be tailored to intensify the search towards the center.

2 Center of the Pareto Front: definition and estimation

Definition. We define the center, \mathbf{C} , of the Pareto front as the projection (intersection in case of a continuous front) of the closest ND point on the Ideal-Nadir line (in the Euclidean objective space). An example of Pareto front center can be seen in Fig. 1. This center corresponds visually to an equilibrium among all objectives. Alternative definitions involving e.g. the barycenter of the Pareto front, are likely to be harder to calculate in high-dimensional spaces. Furthermore, this center has the property of being insensitive to a linear scaling of the objectives in a bi-objective case⁴. \mathbf{C} is also very little sensitive to perturbations of the Ideal or the Nadir point: for a continuous Pareto front, $|\frac{\partial \mathbf{C}_i}{\partial I_i}|$ and $|\frac{\partial \mathbf{C}_i}{\partial N_i}| < 1$, $i = 1, \dots, m$.

Estimation. As the Ideal and the Nadir of the empirical Pareto front will sometimes be weak substitutes for the real ones (leading to a biased estimated center), those two points have to be truly estimated for the purpose of computing the center. The probabilistic nature of the metamodels (GPs) allows to simulate possible responses of the objective functions. Conditional GP simulations are thus performed to create possible Pareto fronts, each of which defines a sample for \mathbf{I} and \mathbf{N} . The estimated Ideal and Nadir are the medians of the samples. The intersection between the line $\hat{\mathcal{L}}$ joining those points and the extended empirical Pareto front is the estimated Pareto front center $\hat{\mathbf{C}}$.

3 Targeting Infill Criteria for Bayesian Optimization

Articulating preferences has already been addressed in evolutionary multi-objective optimization [2]. In Bayesian multi-objective optimization, fitted to costly objectives, new points are sequentially added using an infill criterion whose purpose is to guide the search towards the Pareto front. After having been evaluated, that

⁴ Non-sensitivity to a linear scaling of the objectives is true when the Pareto front intersects the Ideal-Nadir line. Without intersection, exceptions may occur for $m \geq 3$.

point is used to update the metamodel. The Expected Hypervolume Improvement (EHI, [3]) is a commonly used multi-objective infill criterion. EHI chooses the input \mathbf{x} that maximizes the expected growth of the hypervolume dominated by $\widehat{\mathcal{P}}_{\mathbf{y}}$ up to a reference point \mathbf{R} . Classically, \mathbf{R} is taken beyond the observed Nadir, e.g. [4], in order to cover the entire front. WHI [5] is a variant of EHI for targeting particular regions of the objective space through a user-defined weighting function.

Our approach targets the central part of the Pareto front with EHI by solely controlling the reference point \mathbf{R} . Indeed, the choice of \mathbf{R} is instrumental in deciding the combination of objectives for which *Improvement* occurs: $\mathcal{I}_{\mathbf{R}} = \{\mathbf{y} \in Y : \mathbf{y} \preceq \mathbf{R}\}$. Positioning \mathbf{R} at the (estimated) center of the Pareto front $\widehat{\mathbf{C}}$ will favour EHI's search at objective vectors belonging to $\mathcal{I}_{\widehat{\mathbf{C}}}$.

When the objectives are modeled by independent GPs and the used reference point $\mathbf{R} \notin \widehat{\mathcal{P}}_{\mathbf{y}}$, one has $\text{EHI}(\cdot; \mathbf{R}) = \text{mEI}(\cdot; \mathbf{R})$, where mEI stands for the product of the famous mono-objective Expected Improvement (EI) [1] considering \mathbf{R} as the observed minimum in each objective, $\text{mEI}(\mathbf{x}; \mathbf{R}) = \prod_{i=1}^m \text{EI}_i(\mathbf{x}; R_i)$. This observation is particularly appealing from a computational point of view: EHI requires the computation of m -dimensional hypervolumes involving expensive Monte-Carlo estimations in a many-objectives case. For this reason, in our algorithm where $\mathbf{R} \equiv \widehat{\mathbf{C}}$ is not dominated, EHI is replaced by mEI.

Targeting a particular (here, central) part of the Pareto front leads to a fast local convergence. Once \mathbf{R} is on the real Pareto front, the algorithm will try to improve non-improvable values (see left of Fig. 2). To avoid wasting costly evaluations, the convergence has to be checked. To this aim, we estimate $p(\mathbf{y})$, the probability of dominating $\mathbf{y} \in Y$, simulating Pareto fronts through conditional GPs. $p(\mathbf{y})(1 - p(\mathbf{y}))$ is a measure of domination uncertainty, which tends to 0 as $p(\mathbf{y})$ tends to 0 or 1. We assume local convergence when the *line-uncertainty*, $\int_{\widehat{\mathcal{C}}} p(\mathbf{y})(1 - p(\mathbf{y})) d\mathbf{y}$, is small enough. The complete algorithm, called C-EHI, is summarized below. Fig. 1 shows a convergence of $\mathbf{R} \equiv \widehat{\mathbf{C}}$ to the true center of the Pareto front, \mathbf{C} .

Center-Targeting Algorithm (C-EHI)

Create an initial DoE and initialize m GPs

for each objective $f_i, i = 1, \dots, m;$

$t = 0; \text{line-uncertainty} = +\infty;$

while ($\text{line-uncertainty} > \varepsilon$) **and**

($t \leq \text{budget}$) **do**

estimate \mathbf{I}, \mathbf{N} and $\mathbf{C};$

$\mathbf{x}^{(t+1)} = \arg \max_{\mathbf{x} \in X} \text{mEI}(\mathbf{x}; \mathbf{C});$

evaluate $f_i(\mathbf{x}^{(t+1)})$ and update the GPs;

compute $\text{line-uncertainty}(\text{GPs}, \mathbf{I}, \mathbf{N});$

$t = t + 1;$

end

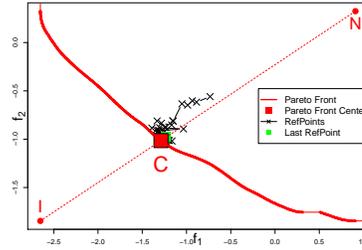


Fig. 1. Reference points \mathbf{R} successively used for directing the search. They lie close to the dashed Ideal-Nadir line (\mathbf{IN}) and lead the algorithm to the center of the Pareto front (\mathbf{C}).

4 Optimization Results

We apply the proposed methodology to a benchmark built from real-world airfoil aerodynamic data. Fig. 2 shows that, compared with standard techniques, the proposed methodology leads to a faster and a more precise convergence to the central part of the Pareto front at the cost of a narrower covering of the front. The results are shown at the iteration which triggers the convergence criterion: only marginal gains would indeed be obtained continuing targeting the same region.

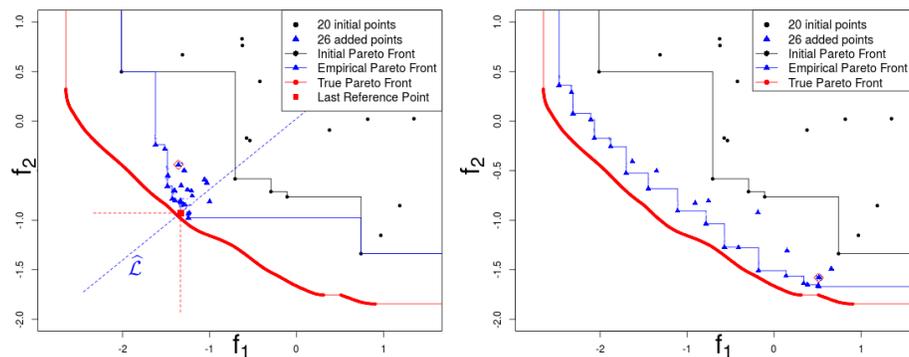


Fig. 2. Two objectives optimization with C-EHI (left). The initial approximation (black) has mainly been improved around the center. Compared with a standard EHI (right), the proposed methodology achieves convergence to the central part of the front. EHI considers more compromises between objectives, but cannot converge within the given budget (46 evaluations).

Further work will consider the continuation of the search to a broader but central part of the Pareto front. That newly targeted region, controlled again through \mathbf{R} , will be as large as possible while being still attainable within the remaining budget.

References

1. Jones, D.R., Schonlau, M., Welch, W.J.: Efficient global optimization of expensive black-box functions. *Journal of Global optimization* **13**(4) (1998) 455–492
2. Bechikh, S., Kessentini, M., Said, L.B., Ghédira, K.: Chap. 4: Preference incorporation in evolutionary multiobjective optimization: A survey of the state-of-the-art. *Advances in Computers* **98** (2015) 141–207
3. Emmerich, M.T., Deutz, A.H., Klinkenberg, J.W.: Hypervolume-based expected improvement: Monotonicity properties and exact computation. In: *Evolutionary Computation (CEC), 2011 IEEE Congress on, IEEE* (2011) 2147–2154
4. Ponweiser, W., Wagner, T., Biermann, D., Vincze, M.: Multiobjective optimization on a limited budget of evaluations using model-assisted s-metric selection. In: *International Conf. on Parallel Problem Solving from Nature, Springer* (2008) 784–794
5. Brockhoff, D., Bader, J., Thiele, L., Zitzler, E.: Directed multiobjective optimization based on the weighted hypervolume indicator. *Journal of Multi-Criteria Decision Analysis* **20**(5-6) (2013) 291–317